

27 DEC 2004

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date
12 February 2004 (12.02.2004)

PCT

(10) International Publication Number
WO 2004/013777 A1

(51) International Patent Classification:
G06F 17/30

(21) International Application Number:
PCT/US2003/025271

(22) International Filing Date: 1 August 2003 (01.08.2003)

English

(25) Filing Language:

English

(26) Publication Language:

(30) Priority Data:
60/401,457

5 August 2002 (05.08.2002) US

(71) Applicant (for all designated States except US): FISH,
Robert [US/US]; 611 Anton Blvd., Suite 1400, Costa
Mesa, CA 92626 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): PONTIUS, Ben
[CA/CA]; 1055 Lucknow St., #802, Halifax, Nova Scotia
B3H 2T3 (CA).

(74) Agents: RUTAN & TUCKER LLP et al.; 611 Anton
Blvd., Suite 1400, Costa Mesa, CA 92626 (US).

(81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

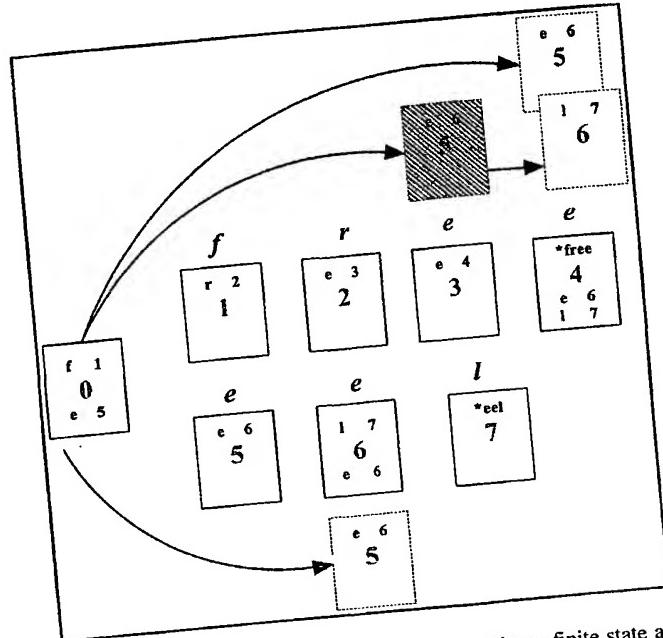
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- with amended claims

Date of publication of the amended claims: 26 August 2004
[Continued on next page]

(54) Title: SYSTEM AND METHOD OF PARALLEL PATTERN MATCHING



WO 2004/013777 A1

(57) Abstract: The present invention provides systems and methods for creating a finite state automata (FSA) (figure 1, blocks 110-180) that matches patterns in parallel including the steps of creating states of the automata from a set of patterns to be matched (figure 2, blocks 210-280) and passing over the patterns a second time adding transitions to the states to match all the possible patterns that can start within the pattern (figure 3, blocks 0-7).



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

AMENDED CLAIMS

[received by the International Bureau on 19 January 2004 (19.01.04);
claims 1-3, 5,6 unchanged]

CLAIMS

What is claimed is:

1. A method for creating finite state automata (FSA) that match patterns in parallel, comprising:
creating states of the finite state automata from a set of patterns to be matched;
passing over the set of patterns a second time; and
adding transitions to the states to match all possible patterns that can start within the set of patterns to be matched.
2. The method of claim 1 further comprising:
iterating through the states;
determining whether input causes a move to an initial state; and
if the initial state has a different move on the input, changing a current state's transition to mirror that of the initial state.
3. A method of creating a FSA that uses array-based transitions for an alphabet of size N, comprising:
representing each state as an object containing an array of N pointers to possible successive states;
using a numeric value of each member of the alphabet as an offset into the array to point to a next state.
4. A method of creating a case-insensitive FSA by making each pattern all one case, comprising:
creating the FSA; and
adding corresponding transitions on each alphabetic character such that testing and conversion of case is not required.
5. A method for matching patterns, comprising:
using a numeric value of less than a complete set of bits of an input as an offset into an array, thereby reducing a size of the array.

6. The method of claim 5 comprising a further step of using a hash function for matching patterns composed of a 128 or 256 alphabet without overhead of larger arrays.